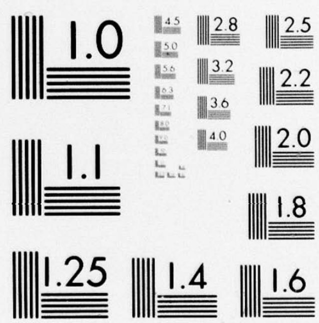
AD-A081 512 MASSACHUSETTS UNIV AMHERST DEPT OF MATHEMATICS AND S--ETC F/G 12/1
NONPARAMETRIC ESTIMATION OF A BIVARIATE SURVIVORSHIP FUNCTION W--ETC(U)
JAN 80 R M KORWAR F49620-79-C-0105
UNCLASSIFIED TR-7 AFOSR-TR-80-0126 NL

| OF |

AD A081512

END
DATE
FILMED
4-80
DTIC


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

18 AFOSR-TR-80-0126 19

LEVEL II

12

SC

6

NONPARAMETRIC ESTIMATION OF A BIVARIATE SURVIVORSHIP FUNCTION WITH DOUBLY CENSORED DATA.

by

10 Ramesh M. Korwar[1]

AFOSR Technical Report, [redacted] 9

14 TR-7

12 13

11 Jan[redacted]80

The University of Massachusetts
Department of Mathematics and Statistics
Amherst, Massachusetts

15

[1]Research sponsored by the Air Force Office of Scientific Research, AFSC, USAF, under Contract AFOSR F49620-79-C-0105

411377 Gw

Nonparametric Estimation of a Bivariate Survivorship Function with Doubly Censored Data

by

Ramesh M. Korwar
University of Massachusetts

Abstract

The problem of nonparametric estimation of a bivariate survivorship function with doubly censored data is considered. A self-consistent estimator is developed by first "reducing" the problem to that of estimation in a singly and right censored situation. This estimator is shown to satisfy a likelihood equation, and its uniqueness is investigated. The results obtained naturally parallel those obtained ~~by Campbell (1979)~~ in the singly censored case.

AMS 1970 subject clasification. Primary 62G05, Secondary 62H99.

Key words and phrases. Bivariate survivorship function, doubly censored data, nonparametric maximum likelihood estimation, self-consistent estimator.

Nonparametric Estimation of a Bivariate
Survivorship Function with Doubly
Censored Data

1. <u>Introduction and summary</u>. In this article we consider the nonparametric estimation of a bivariate survivorship function when data may be both left and right censored. Specifically, let (T_{1j}, T_{2j}), $j = 1,\ldots,n$, be n pairs of true lifelengths where some of the T_{ij}'s are left censored and some right censored. Thus, not all (T_{1j}, T_{2j}) are exactly observable. For each j $(1 \leq j \leq n)$ and for $i = 1,2$, we assume that there are limits of observation L_{ij} and $U_{ij} (L_{ij} \leq U_{ij})$ which are either random variables independent of each other and (T_{1j}, T_{2j}) or are fixed constants. The recorded information is

$$X_{ij} = \max[\min(T_{ij}, U_{ij}), L_{ij}], \quad i = 1,2,\ j = 1,\ldots,n \ .$$

Also, for each $j = 1,\ldots,n$ and $i = 1,2$, it is known whether $X_{ij} = L_{ij}$ (i.e. $T_{ij} \leq L_{ij}$ and T_{ij} is left censored or a "late entry"), or $X_{ij} = U_{ij}$ (i.e. $T_{ij} > U_{ij}$ and T_{ij} is right censored or a "loss"), or $X_{ij} = T_{ij}$ (i.e. $L_{ij} < T_{ij} \leq U_{ij}$ and T_{ij} is a "death").

Recently Campbell (1979) considered the problem when observations are randomly and right censored. He developed a "self-consistent" (SC) nonparametric estimator and showed, among other things, that the SC satisfied a nonparametric likelihood equation and was unique up to the final censored values in any dimension. A similar SC estimator was developed and studied by Korwar and Dahiya (1979). Here we extend Campbell's results to the

This research was supported by a research contract from the U.S. Air Force Office of Scientific Research, AFOSR, under Contract #F49620-79-C-0105.

situation of doubly censored data. Earlier, Turnbull (1974) considered the univariate situation, and used an ingenious approach to "reduce" the doubly censored situation to the right censored one for which the Kaplan-Meir (KL) (1958) estimator was readily applicable. In extending Campbell's results here we use the very same approach of Turnbull, though modified, of "reducing" the doubly censored situation to that of the right and singly censored one.

Assume, then, that there is, for each component of the bivariate variable, a natural discrete time $t_{i0} = 0 < t_{i1} < \ldots < t_{im_i}$, $i = 1,2$. (This is an appropriate assumption to make if observations are made only at discrete times.) Alternatively, we might assume that they are grouped and for each rectangle $(t_{1i-1}, t_{1i}] \times (t_{2j-1}, t_{2j}]$ we have the following information available:

δ_{ij} = # of pairs (T_{1k}, T_{2k}) such that $t_{1i-1} < T_{1k} \le t_{1,i}$ and $t_{2j-1} < T_{2k} \le t_{2j}$ ("double deaths");

α_{ij} = # of pairs (T_{1k}, T_{2k}) such that $T_{1k} > T_{1i}$ and $t_{2j-1} < T_{2k} \le t_{2j}$ ("loss in the first component and death in the second");

β_{ij} = # of pairs (T_{1k}, T_{2k}) such that $t_{1i-1} < T_{1k} \le t_{1i}$ and $T_{2k} > t_{2j}$ ("death in the first component and loss in the second");

λ_{ij} = # of pairs (T_{1k}, T_{2k}) such that $T_{1k} > t_{1i}$ and $T_{2k} > t_{2j}$ ("double losses");

μ^1_{ij} = # of pairs (T_{1k}, T_{2k}) such that $T_{1k} \le t_{1i}$ and $T_{2k} \le t_{2j}$ ("double late entries");

μ^2_{ij} = # of pairs (T_{1k}, T_{2k}) such that $T_{1k} > t_{1i}$ and $T_{2k} \le t_{2j}$ ("loss in the first component and late entry for the second");

μ^3_{ij} = # of pairs (T_{1k}, T_{2k}) such that $T_{1k} \le t_{1i}$ and $T_{2k} > t_{2j}$ ("late entry for the first component and loss in the second");

μ^4_{ij} = # of pairs (T_{1k}, T_{2k}) such that $t_{1i-1} < T_{1k} \le t_{1i}$ and $T_{2k} \le t_{2j}$ ("death in the first component and late entry in the second"); and

μ^5_{ij} = # of pairs (T_{1k}, T_{2k}) such that $T_{1k} \le t_{1i}$ and $t_{2j-1} < T_{2k} \le t_{2j}$ ("late entry for the first component and death in the second").

As in Turnbull (1974), we assume that each of the late entries μ^3_{ij} or μ^5_{ij}, (μ^2_{ij}, or μ^4_{ij}) in the first (second) component occurs at the end of the interval $(t_{1i-1}, t_{1i}]((t_{2j-1}, t_{2j}])$; that each of the double late entries μ^1_{ij} (double losses $\lambda_{i-1,j-1}$) occurs at the right hand top (left hand botton) corner of the rectangle $(t_{1i-1}, t_{1i}] \times (t_{2j-1}, t_{2j}]$. Finally, we also assume that each of the α_{ij} or μ^2_{ij} (β_{ij} or μ^3_{ij}) losses in the first (second) component occurs at the beginning of the interval $(t_{1i-1}, t_{1i}]$ $((t_{2j-1}, t_{2j}])$. Alternative assumptions are also possible which we do not pursue here.

We desire to estimate $F_{ij} = P[T_1 > t_{1i}, T_2 > t_{2j}]$, $i = 0,\ldots,m_1$, $j = 0,\ldots,m_2$.

2. The Self-consistent estimator. We first reduce the problem to that in which there is only right censoring. To this end, consider first the μ^1_{ij} double late entries at (t_{1i}, t_{2j}). Of these, the expected number of double

deaths in the rectangle $(t_{1,i'-1}, t_{1i'}] \times (t_{2,j'-1}, t_{2j'}]$, for $i' \leq i$, $j' \leq j$, is estimated by $\mu^1_{ij}\, k^1_{ij}$, where k^1_{ij} is an estimate of

$P[t_{1,i'-1} < T_1 \leq t_{1i'}, t_{2,j'-1} < T_2 \leq t_{2j} \mid T_1 \leq t_{1i}, T_2 \leq t_{2j}] = \Delta_{i'j'} / \rho_{ij}$,

where $\Delta_{i'j'} = F_{i'-1,j'-1} - F_{i'-1,j'} - F_{i',j'-1} + F_{i',j'}$, $\rho_{ij} = 1 - F_{0,j} - F_{i,0} + F_{i,j}$, and $F_{ij} = P[T_1 > t_{1i}, T_2 > t_{2j}]$. Similarly, of the $\mu^4_{ij}(\mu^5_{ij})$ single late entries, the expected number of double deaths in the rectangle $(t_{1,i-1}, t_{1i}] \times (t_{2,j'-1}, t_{2j'}]$ is estimated by $\mu^4_{ij}\, k^4_{ij}(\mu^5_{ij}\, k^5_{ij})$ where $k^4_{ij}(k^5_{ij})$ is an estimate of $P[t_{1i-1} < T_1 \leq t_{1i}, t_{2j'-1} < T_2 \leq t_{2,j'} \mid t_{1i-1} < T_1 \leq t_{1i}, T_2 > t_{2j}] = \Delta_{ij'}/(R_{i0} - R_{ij})$ $(P[t_{1,i'-1} < T_1 \leq t_{1i'}, t_{2,j-1} < T_2 \leq t_{2j} \mid T_1 > t_{1i}, t_{2,j-1} < T_2 \leq t_{2j}] = \Delta_{ij}/(Q_{0j} - Q_{ij}))$, where $\Delta_{ij}'(\Delta_{i'j})$ is defined above and $R_{ij} = F_{i-1,j} - F_{ij}(Q_{ij} = F_{ij-1} - F_{ij})$. Of the $\mu^3_{ij}(\mu^2_{ij})$ single late entries, the expected number of deaths in the rectangle $(t_{1,i'-1}, t_{1i'}] \times (t_j, \infty)$ is estimated by $\mu^3_{ij}k^3_{ij}(\mu^2_{ij}k^2_{ij})$ where $k^3_{ij}(k^2_{ij})$ is an estimate of $P[t_{1i'-1} < T_1 \leq t_{1i'}, T_2 > t_j \mid T_1 \leq t_{1i}, T_2 > t_{2j}] = R_{i',j}/(F_{0,j} - F_{i,j})$ $(P[T_1 > t_{1i}, t_{2j'-1} < T_2 \leq t_{2j'} \mid T_1 > t_{1i}, T_2 \leq t_{2j}] = Q_{ij'}/(F_{i,0} - F_{i,j}))$. Now, consider the "reduced" singly, right censored problem with

$$\delta'_{ij} = \delta_{ij} + \sum_{\substack{i'>i \\ j'\geq j}} \mu^1_{ij}\, \hat{\Delta}_{ij}/\hat{\rho}_{i'j'} + \sum_{j'\geq j} \mu^4_{ij'}\, \hat{\Delta}_{ij}/(\hat{R}_{i,0} - \hat{R}_{i,j'}) + \sum_{i'\geq i} \mu^5_{i'j} \hat{\Delta}_{ij}/(Q_{0,j} - \hat{Q}_{i'j}), \tag{2.1}$$

$$\alpha'_{ij} = \alpha_{ij} + \sum_{j'\geq j} \mu^2_{kj'}\, \hat{Q}_{ij}/(\hat{F}_{i,0} - \hat{F}_{i,j'}),\quad \beta'_{ij} = \beta_{ij} + \sum_{i'\geq i} \mu^3_{i'j}\, \hat{R}_{ij}/(\hat{F}_{0j} - \hat{F}_{i'j}),$$

$$\lambda'_{ij} = \lambda_{ij}$$

where the caret over a probability denotes an estimate of the corresponding probability. From Campbell (1979) a nonparametric estimator F'_{ij} of $F_{i,j} = P[T_1 > t_{1i}, T_2 > t_{2j}]$ satisfies the self-consistency equation

$$n\hat{F}'_{ij} = N_{ij} + \sum_{\substack{\ell>j \\ k<i}} \alpha'_{k\ell} \frac{Q'_{i\ell}}{Q'_{k\ell}} + \sum_{\substack{k>i \\ \ell<j}} \beta'_{k\ell} \frac{\hat{R}'_{kj}}{\hat{R}'_{k\ell}} + \sum_{\substack{k<i \\ \text{or } \ell<j}} \lambda'_{k\ell} \frac{\hat{F}'_{\max(i,k),\max(i,\ell)}}{\hat{F}'_{k\ell}} ,$$

(2.2) $$N_{ij} = \sum_{k>i,\ell>j} \delta'_{k\ell} + \sum_{k\geq i,\ell>j} \alpha'_{k\ell} + \sum_{k>i,\ell\geq j} \beta'_{k\ell} + \sum_{k\geq i,\ell\geq j} \lambda'_{k\ell} .$$

Following Efron (1967) we say that the estimates $\{\hat{F}_{ij}\}$ are self-consistent if $\hat{F}'_{ij} = \hat{F}_{ij}$. Thus the self-consistent estimates $\hat{F}_{ij}$ satisfy

(2.3) $$n\hat{F}_{ij} = N_{ij} + \sum_{\substack{\ell>j \\ k<i}} \alpha'_{k\ell} \frac{\hat{Q}_{i\ell}}{\hat{Q}_{k\ell}} + \sum_{\substack{k>i \\ \ell<j}} \beta'_{k\ell} \frac{\hat{R}_{kj}}{\hat{R}_{k\ell}} + \sum_{\substack{k<i \\ \text{or } \ell<j}} \lambda'_{k\ell} \frac{\hat{F}_{\max(ik),\max(j,\ell)}}{\hat{F}_{k\ell}}$$

where $(\alpha'_{k\ell}, \beta'_{k\ell}, \delta'_{k\ell}, \lambda'_{k\ell})$ and N_{ij} are given by (2.1) and (2.2) respectively, and where $\hat{Q}_{ij} = \hat{F}_{i,1-1} - \hat{F}_{ij}$ and $\hat{R}_{ij} = \hat{F}_{i-1,j} - \hat{F}_{ij}$.

An obvious iterative procedure to compute the $\hat{F}_{ij}$'s is as follows:

(1) Start with the empirical distribution function $\{\hat{F}^0_{ij}\}$, $\hat{F}^0_{ij} = N_{ij}/n$, as the initial estimators.

(2) Form $\delta'_{ij}, \alpha'_{ij}, \beta'_{ij}$ and λ'_{ij}, as described above, using $\{\hat{F}^0_{ij}\}$.

(3) Obtain improved estimates $\{\hat{F}^1_{ij}\}$ by using

$$n\hat{F}^1_{ij} = N_{ij} + \sum_{\substack{\ell>j \\ k<i}} \alpha'_{k\ell} \frac{\hat{Q}^0_{i\ell}}{\hat{Q}^0_{k\ell}} + \sum_{\substack{k>\ell \\ \ell<j}} \beta'_{k\ell} \frac{\hat{R}^0_{kj}}{\hat{R}^0_{k\ell}} + \sum_{\substack{k<i \\ \text{or } \ell<j}} \lambda'_{k\ell} \frac{\hat{F}^0_{\max(i,k),\max(i,\ell)}}{\hat{F}^0_{k\ell}}$$

(4) Go to step (2) with $\{\hat{F}^1_{ij}\}$ in place of $\{\hat{F}^0_{ij}\}$ etc.

(5) Stop when the requirements on the accuracy are met. Convergence could be faster if we started with the Campbell estimates $\{\hat{F}^{00}_{ij}\}$ satisfying (2.3) without the dashes.

3. Properties of the SC estimator. In this section we state and prove some important properties of the self-consistent estimates $\{\hat{F}_{ij}\}$ (2.3). These properties naturally parallel properties for Campbell's (1979) estimates for the right censoring only situation.

The likelihood function L is

$$L = \prod_{i=1,j=1}^{m_1,m_2} \Delta_{ij}^{\delta_{ij}} Q_{ij}^{\alpha_{ij}} R_{ij}^{\beta_{ij}} F_{ij}^{\lambda_{ij}} \rho_{ij}^{\mu^1_{ij}} (F_{i0} - F_{ij})^{\mu^2_{ij}} (F_{0,j} - F_{ij})^{\mu^3_{ij}} (R_{i0} - R_{ij})^{\mu^4_{ij}} \cdot (Q_{0j} - Q_{ij})^{\mu^5_{ij}} \tag{3.1}$$

and the likelihood equation obtained by partially differentiating ℓnL with respect to F_{ij} and setting it to zero is

$$\frac{\partial \ell nL}{\partial F_{ij}} = \frac{\delta_{ij}}{\Delta_{ij}} - \frac{\delta_{i,j+1}}{\Delta_{i,j+1}} - \frac{\delta_{i+1,j}}{\Delta_{i+1,j}} + \frac{\delta_{i+1,j+1}}{\Delta_{i+1,j+1}} + \frac{\lambda_{ij}}{F_{ij}} - \frac{\alpha_{ij}}{Q_{ij}} + \frac{\alpha_{i,j+1}}{Q_{i,j+1}} - \frac{\beta_{ij}}{R_{ij}}$$
$$+ \frac{\beta_{i+1,j}}{R_{i+1,j}} + \frac{\mu^1_{ij}}{\rho_{ij}} - \frac{\mu^2_{ij}}{F_{i0}-F_{ij}} - \frac{\mu^3_{ij}}{F_{0j}-F_{ij}} + \frac{\mu^4_{ij}}{R_{i0}-R_{ij}} - \frac{\mu^4_{i+1,j}}{R_{i+1,0}-R_{i+1,j}} + \frac{\mu^5_{ij}}{Q_{0j}-Q_{ij}}$$
$$- \frac{\mu^5_{ij+1}}{Q_{0,j+1}-Q_{i,j+1}} = 0 ,$$

$$\frac{\partial \ell nL}{\partial F_{i0}} = - \frac{\delta_{i1}}{\Delta_{i1}} + \frac{\delta_{i+1,1}}{\Delta_{i+1,1}} + \frac{\alpha_{i1}}{Q_{i1}} - \frac{\mu^1_{ij}}{\rho_{ij}} + \frac{\mu^2_{ij}}{F_{i0}-F_{ij}} - \frac{\mu^4_{ij}}{R_{i0}-R_{ij}} + \frac{\mu^4_{i+1,j}}{R_{i+1,0}-R_{i+1,j}} = 0 ,$$

(3.2)

$$\frac{\partial \ell nL}{\partial F_{0j}} = - \frac{\delta_{1j}}{\Delta_{1j}} + \frac{\delta_{1,j+1}}{\Delta_{1,j+1}} + \frac{\beta_{1j}}{R_{1j}} - \frac{\mu^1_{ij}}{\rho_{ij}} + \frac{\mu^3_{ij}}{F_{0j}-F_{ij}} - \frac{\mu^5_{ij}}{Q_{0j}-Q_{ij}} + \frac{\mu^5_{i,j+1}}{Q_{0,j+1}-Q_{i,j+1}} = 0 .$$

We are now ready to prove

Theorem 1: Any self-consistent solution $\{\hat{F}_{ij}\}$ of (2.3) will also satisfy the likelihood equation (3.2) with $(\hat{F}_{ij},\hat{Q}_{ij},\hat{R}_{ij})$ in place of (F_{ij},Q_{ij},R_{ij}).

Proof: Recast, using (2.1), the likelihood equation (3.2). This is easily seen to be

$$\frac{\delta'_{ij}}{\Delta_{ij}} - \frac{\delta'_{i,j+1}}{\Delta_{i,j+1}} - \frac{\delta'_{i+1,j}}{\Delta_{i+1,j}} + \frac{\delta'_{i+1,j+1}}{\Delta_{i+1,j+1}} + \frac{\lambda'_{ij}}{F_{ij}} - \frac{\alpha'_{ij}}{Q_{ij}} + \frac{\alpha'_{i,j+1}}{Q_{i,j+1}} - \frac{\beta'_{ij}}{R_{ij}} + \frac{\beta'_{i+1,j}}{R_{i+1,j}} = 0 ,$$

$$-\frac{\delta'_{i1}}{\Delta_{i1}} + \frac{\delta'_{i+1,1}}{\Delta_{i+1,1}} + \frac{\alpha'_{i1}}{Q_{i1}} = 0 \quad , \text{ and}$$

$$-\frac{\delta'_{1j}}{\Delta_{1j}} + \frac{\delta'_{1,j+1}}{\Delta_{1,j+1}} + \frac{\beta'_{1j}}{R_{1j}} = 0 \quad .$$

The rest follows from Theorem 2 of Campbell (1979). □

Turning to the existence and uniqueness of a solution of (2.3), one might ask whether there exists at all a solution $\{\hat{F}_{ij}\}$ of (2.3) and if so is it unique? The difficulties are similar to those encountered in the singly and right censored case. For a discussion and a conjecture about existence, see Campbell (1979). The following theorem is aimed at the uniqueness question:

Theorem 2: The matrix $\{-\frac{\partial^2 nL}{\partial F_{ij}\partial F_{k\ell}}\}$ is nonnegative definite.

Proof: Write the likelihood (3.1) as

$$L = L_1 L_2$$

where

$$L_1 = \prod_{i,j} \Delta_{ij}^{\delta_{ij}} F_{ij}^{\lambda_{ij}} Q_{ij}^{\alpha_{ij}} R_{ij}^{\beta_{ij}} ,$$

and

$$(3.3) \qquad L_2 = \prod_{i,j} \rho_{ij}^{\mu_{ij}^1} (F_{i0}-F_{ij})^{\mu_{ij}^2} (F_{0j}-F_{ij})^{\mu_{ij}^3} (R_{i0}-R_{ij})^{\mu_{ij}^4} (Q_{0j}-Q_{ij})^{\mu_{ij}^5} .$$

Now, Campbell (1979) has shown (his Theorem 3) that the matrix $\{-\frac{\partial^2 \ell n L_1}{\partial F_{ij} \partial F_{k\ell}}\}$ is non-negative definite. Thus it suffices to show that $\{-\frac{\partial^2 \ell n L_2}{\partial F_{ij} \partial F_{k\ell}}\}$ also is. From (3.3) and the definitions of ρ_{ij}, Q_{ij} and R_{ij}, letting $-\frac{\partial^2 \ell n L_2}{\partial F_{ij} \partial F_{k\ell}} = M_{(ij)(k\ell)}$, we find that

$$M_{(ij)(ij)} = m_{ij}^1 + m_{ij}^2 + m_{ij}^3 + m_{ij}^4 + m_{i+1,j}^4 + m_{ij}^5 + m_{i,j+1}^5 ,$$

$$M_{(ij)(i-1,j)} = - m_{ij}^4 , \ M_{(ij)(i,j-1)} = - m_{ij}^5 , \ M_{(ij)(k\ell)} = 0 \ \text{ if } |i-k|+|j-\ell| \geq 2,$$

$$M_{(ij)(i-1,0)} = m_{ij}^4 , \ M_{(ij)(i0)} = -(m_{ij}^1 + m_{ij}^2 + m_{ij}^4 + m_{i+1,j}^4), \ M_{(ij)(i+1,0)} = m_{i+1,j}^4 ,$$

$$M_{(ij)(0,j-1)} = m_{ij}^5 , \ M_{(ij)(0j)} = -(m_{ij}^1 + m_{ij}^3 + m_{ij}^5 + m_{i,j+1}^5), \ M_{(ij)(0,j+1)} = m_{i,j+1}^5 ,$$

$$M_{(ij)(k0)} = 0 \ \text{ if } |i-k| \geq 2, \ M_{(ij)(0\ell)} = 0 \ \text{ if } |j-\ell| \geq 2 ,$$

$$M_{(0j)(0j)} = m_{ij}^1 + m_{ij}^3 + m_{ij}^5 + m_{i,j+1}^5 , \ M_{(0j)(0,j-1)} = - m_{ij}^5 ,$$

$$M_{(0j)(i0)} = m_{ij}^1 , \ M_{(0j)(0\ell)} = 0 \ \text{ if } |\ell-j| \geq 2,$$

$$M_{(i0)(i0)} = m_{ij}^1 + m_{ij}^2 + m_{ij}^4 + m_{i+1,j}^4 , \ M_{(i0)(i-1,0)} = - m_{ij}^4 , \ M_{(i0)(k0)} = 0 \ \text{ if } |k-i| \geq 2$$

where

$$m_{ij}^1 = \mu_{ij}^1/\rho_{ij}^2,\ m_{ij}^2 = \mu_{ij}^2/(F_{i0} - F_{ij})^2,\ m_{ij}^3 = \mu_{ij}^3/(F_{0j} - F_{ij})^2 \quad ,$$

$$m_{ij}^4 = \mu_{ij}^4/(R_{i0} - R_{ij})^2,\ m_{ij}^5 = \mu_{ij}^5/(Q_{0j} - Q_{ij})^2 \quad .$$

Note that M(as well as the matrices to follow) is $m \times m$, where $m = m_1m_2 + m_1 + m_2$ Now decompose M as follows: Let A_{ij} be the matrix with 1's at (0j)(0j), (0j)(i0), (i0)(0j), (i0)(i0), (ij)(ij), - 1's at (0j)(ij),(i0)(ij),(ij)(i0), and 0's everywhere else. Let A_{ij}^2 be the matrix with 1's at (i0)(i0), (ij)(ij), -1's at (i0)(ij), (ij)(i0) and 0's everywhere else. Let A_{ij}^3 be the Matrix with 1's at (0j)(0j), (ij)(ij), - 1's at (0j)(ij), (ij,0j), and 0's everywhere else. Let A_{ij}^5 be the matrix with 1's at (0,j-1)(0,j-1), (0,j-1)(ij), (0j)((0j),(0j)(i,j-1), (i,j-1)(0j), (i,j-1)(i,j-1), (0,j-1)(ij), (ij)(ij), -1's at (0,j-1)(0j), (0,j-1)(i,j-1),(0j)(0,j-1), (0,j)(ij), (i,j-1)(0,j-1), (i,j-1)(ij), (ij)(0j), (ij)(i,j-1), and 0's everywhere else. Finally, let A_{ij}^4 be the matrix with 1's at (i-1,0)(i-1,0), (i-1,0)(i,j), (i0)(i0), (i0)(i-1,j), (i-1,j)(i0), (i-1,j)(i-1,j), (ij)(i-1,0),(ij)(ij), - 1's at (i-1,0)(i0), (i-1,0)(i-1,j), (i0)(i-1,0), (i0)(ij), (i-1,j)(i-1,0), (i-1,j)(ij), (ij)(i0), (ij)(i-1,j), and 0's everywhere else. It is easily verified that $M = \sum_{j=1}^{m_2} \sum_{i=1}^{m_1} \sum_{k=1}^{5} m_{ij}^k A_{ij}^k$. Since μ_{ij}^k's are nonnegative numbers, the proof will be complete if we show that each A_{ij}^k is nonnegative definite. A_{ij}^2 and A_{ij}^3 are obviously nonnegative definite. The highest order nontrivial principal minor of A_{ij}^1 is a third order determinant with two identical rows, all the second order principal minors of this determinant vanish; and the main diagonal has 1's on it. Hence A_{ij}^1 is also nonnegative definite.

Finally, the highest order nontrivial principal minor of $A_{ij}^{4}(A_{ij}^{5})$ is a fourth order determinant which has two pairs of identical rows. Hence, itself and all its principal minors of order three vanish. The second order principal minors of this determinant all vanish, and the entries on the main diagonal are all 1's . □

Theorem 2 implies that the likelihood L (3.1) is convex in the F_{ij}'s.

References

[1] Campbell, G. (1979). Nonparametric bivariate estimation with randomly censored data. Purdue University, Department of Statistics, Mimeoseries #79-25.

[2] Efron, B. (1967). The two sample problem with censored data. Proc. Fifth Berkeley Symp. Math Statist & Prob. 4, 831-53.

[3] Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-81.

[4] Korwar, R.M. and Dahiya, R.C. (1979). Estimation of a bivariate distribution function from incomplete observations. University of Massachusetts, Department of Mathematics and Statistics, AFOSR Tech. Report #2.

[5] Turnbull, B.W. (1974). Nonparametric estimation of a survivorship function with doubly censored data. J. Amer. Statist. Assoc. 69, 169-73.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR Report No. 7	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) NONPARAMETRIC ESTIMATION OF A BIVARIATE SURVIVORSHIP FUNCTION WITH DOUBLY CENSORED DATA		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ramesh M. Korwar		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS The University of Massachusetts Department of Mathematics & Statistics Amherst, MA 01003		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Bolling Air Force Base, Washington DC 20332		12. REPORT DATE January 1980
		13. NUMBER OF PAGES 11
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bivariate survivorship function
Doubly censored data
Nonparametric maximum likelihood estimation
Self-consistent estimator

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The problem of nonparametric estimation of a bivariate survivorship function with doubly censored data is considered. A self-consistent estimator is developed by first "reducing" the problem to that of estimation in a singly and right censored situation. This estimator is shown to satisfy a likelihood equation, and its uniqueness is investigated. The results obtained naturally parallel those obtained by Campbell (1979) in the singly censored case.

DD FORM 1473

END

DATE
FILMED
4-80

DTIC